

Fundamental limits on the accuracy of optical phase estimation from rate-distortion theory

Ranjith Nair

*Department of Electrical & Computer Engineering,
National University of Singapore, Singapore 117583*

(Dated: April 17, 2012)

Lower bounds are derived on the average mean-squared error of optical phase estimation in a Bayesian framework using classical rate-distortion theory in conjunction with the classical capacity of the lossy and lossless optical channel under phase modulation. With no optical loss, the bound displays Heisenberg-limit scaling of the mean-squared error $\delta\Phi^2 \sim 1/N_S^2$ where N_S is the average number of photons in the probe state. In the presence of nonzero loss, a lower bound with standard-quantum-limit (SQL) asymptotic scaling is derived. The bounds themselves are non-asymptotic and valid for any prior probability distribution of the phase.

PACS numbers: 42.50.Ex, 42.50.St, 06.20.-f, 03.67.Hk

I. INTRODUCTION

It has long been argued that the mean-squared error in sensing an optical phase shift can exhibit, at best, an inverse quadratic scaling with the mean number of photons in the quantum state used to sense the phase [1]. Recently, following some claims that this so-called “Heisenberg limit” (or H limit) on phase estimation may be beaten [2–4], several authors have revived the subject by providing rigorous proofs of lower bounds, not limited to optical interferometry, with H limit scaling [5–9]. Interestingly, these proofs use diverse techniques, namely, the speed limit on quantum evolutions [5, 8], the entropic uncertainty relations [6, 9], and the quantum Ziv-Zakai bound [7, 8]. In this paper, we develop another technique for obtaining lower bounds – classical rate-distortion theory – that was introduced into quantum metrology in [10], and apply it to both lossless and lossy optical phase estimation.

Rate-distortion theory [11] is a branch of classical information theory [12, 13] that was introduced in ref. [12] and elaborated in ref. [14] by Shannon, and forms the theoretical basis for the lossy compression of data sources. In the simplest scenario involving a continuous data source, the source generates an output modeled as a random variable U , and we wish to map U to another random variable V – one that perhaps presents lesser storage requirements – in such a way that a predefined distortion measure d such as the average mean-squared error between U and V is kept below a tolerable level d^* . Roughly speaking, rate-distortion theory tells us how much information must remain in V in order to do so, and shows that coded schemes can achieve this compression limit. A fascinating historical introduction into the theory and practice of lossy data compression may be found in ref. [15].

The development of quantum information theory [16, 17] in the past few decades has been much influenced by the ideas of classical information theory, including rate-distortion theory. One of the first results of quantum information theory, the noiseless coding theorem of Schumacher [18] is a quantum version of Shannon’s noiseless

source coding theorem, which itself corresponds to the rate-distortion theory with allowed distortion set to zero. More recently [19], there have been efforts to formulate a quantum rate-distortion theory applicable to the lossy compression of quantum rather than classical information sources.

In this paper, we apply classical rate-distortion theory to the problem of estimation of an optical phase using quantum probe states whose mean energy, i.e., photon number, is constrained. We adopt a Bayesian approach in which an arbitrary prior distribution of the phase is assumed and the squared error is averaged over this distribution. We first consider the ideal lossless case in which we obtain a rigorous lower bound for the mean-squared error that exhibits H limit scaling. This derivation applies to any multimode probe state as long as the energy in the modes sensing the phase is constrained to be at most N_S . We then allow for the presence of loss, and show that, for any probe state with a single mode undergoing the phase shift, the mean-squared phase error scales inversely as the mean number of photons in the probe state, i.e., it exhibits the “Standard Quantum Limit” (SQL) scaling of coherent states [1].

II. RATE-DISTORTION THEORY AND THE INFORMATION TRANSMISSION INEQUALITY

Consider the random variable Φ representing the phase shift to be sensed, and let its prior probability density be $P_\Phi(\phi)$. In rate-distortion theory, Φ is viewed as a data source with differential entropy $h(\Phi)$ given by

$$h(\Phi) = - \int_0^{2\pi} d\phi P_\Phi(\phi) \ln P_\Phi(\phi) \quad (1)$$

and measured in nats/symbol. For another random variable $\hat{\Phi}$ representing an estimate of Φ , we can define the

mean-squared error (or MSE) as

$$\delta\Phi^2 := \mathbb{E} [\hat{\Phi} - \Phi]^2 = \int_0^{2\pi} P_{\Phi}(\phi) P_{\hat{\Phi}|\Phi}(\hat{\phi}|\phi) (\hat{\phi} - \phi)^2, \quad (2)$$

where $P_{\hat{\Phi}|\Phi}(\hat{\phi}|\phi)$ is the conditional density of $\hat{\Phi}$ given Φ .

The mean-squared error (2) is an example of a *distortion measure*, denoted $d(\Phi, \hat{\Phi})$ in general, which is an ensemble average of a numerical function of the source output and the estimate that measures how far they are different for the purposes of a particular application. A variety of distortion measures may be used [13], but only the mean-squared error distortion measure is considered in this paper. The *rate-distortion function* $R(D)$ is defined as [13, 14]

$$R(D) = \inf_{P_{\hat{\Phi}|\Phi}(\hat{\phi}|\phi): d(\Phi, \hat{\Phi}) \leq D} I(\Phi; \hat{\Phi}), \quad (3)$$

where the quantity being minimized is the mutual information $I(\Phi; \hat{\Phi})$ between the source and estimate and the infimum is over all conditional distributions $P_{\hat{\Phi}|\Phi}(\hat{\phi}|\phi)$ that yield average distortion less than or equal to D . Note that $R(D)$ depends on the prior distribution $P_{\Phi}(\phi)$ and the distortion measure $d(\Phi, \hat{\Phi})$.

The rate-distortion function $R(D)$ may be thought of informally as the amount of non-redundant information per symbol emitted by the source, given that we allow for a distortion of up to D in a reconstructed version of the source. Examples of the computation of $R(D)$ for some standard sources and distortion measures may be found in [13, 14], although numerical evaluation is usually required for an arbitrary source. In general, the function $R(D)$ and its inverse $D(R)$ are decreasing and convex in their respective arguments. For the mean-squared error distortion measure, the following lower bound on $R(D)$ may be used [14, 15]:-

$$R(D) \geq \frac{1}{2} \ln \left(\frac{Q_{\Phi}}{D} \right) \equiv \underline{R}(D), \quad (4)$$

where

$$Q_{\Phi} = \frac{1}{2\pi e} e^{2h(\Phi)}, \quad (5)$$

is the *entropy power* of Φ [13]. Note that the bound is useful for $D \in (0, Q_{\Phi}]$ (outside which it can be replaced with zero) and is convex and decreasing on this interval. The following lower bound on $D(R)$ follows from eq. (4):-

$$D(R) \geq Q_{\Phi} e^{-2R} \equiv \underline{D}(R). \quad (6)$$

The operational significance of the rate-distortion function is elucidated by the positive and converse parts of the noisy source coding theorems of Shannon [14]. For our purpose of obtaining lower bounds on the achievable distortion, the converse part is of primary relevance. The fundamental result, called the *Information Transmission Inequality* in [15], is stated below (Refer Fig. 1).

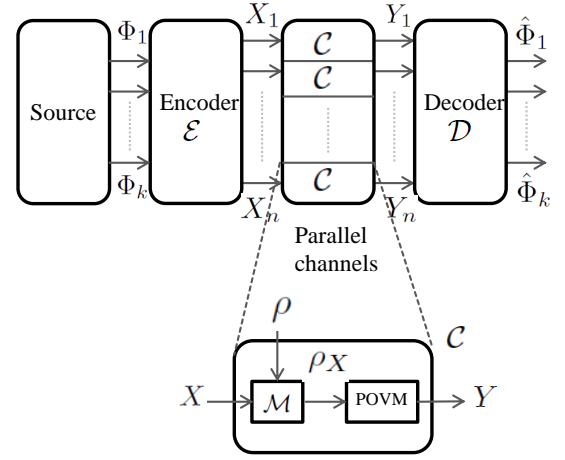


FIG. 1. Block diagram of the estimation scenario to which the Information Transmission Inequality applies. The blow-up shows how each of the parallel channels \mathcal{C} is realized by a modulation of X into density operators of a quantum system followed by a POVM measurement on the system.

Theorem 1 (Information Transmission Inequality – Theorem 1 of ref. [14]). *Given k independent and identically distributed (i.i.d.) source outputs $\Phi = \Phi_1, \dots, \Phi_k$, each with prior distribution $P_{\Phi}(\phi)$. For a given distortion measure $d(\Phi, \hat{\Phi})$, let the rate-distortion function of the source be $R(D)$ nats/symbol. Given an encoder \mathcal{E} that maps $\Phi = \Phi_1, \dots, \Phi_k$ to an n -symbol-long codeword $\mathbf{X} = X_1, \dots, X_n$ that is transmitted over a channel \mathcal{C} with capacity C nats/use. Let the channel output codeword be $\mathbf{Y} = Y_1, \dots, Y_n$ which is mapped by a decoder \mathcal{D} to an estimate $\hat{\Phi} = \hat{\Phi}_1, \dots, \hat{\Phi}_k$ of Φ . Defining a per-symbol average distortion measure $\bar{d}(\Phi, \hat{\Phi}) = \sum_{i=1}^k d(\Phi_i, \hat{\Phi}_i)/k$, we have*

$$\bar{d}(\Phi, \hat{\Phi}) \geq D \left(\frac{n}{k} C \right), \quad (7)$$

where $D(\cdot)$ is the function inverse to $R(D)$.

Note that, unlike the positive part of the noisy source coding theorem which applies in the asymptotic limit of long codes with $n \rightarrow \infty$, Theorem 1 applies to any given system of the form of Fig. 1. The application of Theorem 1 to obtain performance lower bounds in classical estimation and communication is well-known [20]. Its application to quantum metrology was first proposed in [10] and is made as follows. The key point is to implement the classical channel \mathcal{C} appearing in Fig. 1 using a quantum system in the following way. Given a codeword symbol X , we implement a *modulation* map \mathcal{M} that takes the symbol X and a given *probe state* ρ in the Hilbert space \mathcal{H} of the quantum system of interest into another density operator $\rho_X \in \mathcal{H}$. We then make a *measurement* on ρ_X described by a Positive-Operator-Valued Measure (POVM) $\{\Pi_Y\}$ [16], whose outcome Y is the output codeword symbol (see inset to Fig. 1). Any such choice of

probe state, modulation map, and POVM measurement induces a probability transition matrix $P_{Y|X}(y|x)$, i.e., a classical channel, for which a channel capacity C may be defined. This C may then be used in eq. (7) to yield a lower bound on the distortion. The calculation of C can be made to incorporate any constraints relevant to the sensing problem, e.g., an energy constraint on the probe state, a constraint on the kind of modulation allowed, or a constraint on the measurement POVM.

In the context of quantum optics, we can consider a fixed class of probe states, e.g., coherent or quadrature squeezed states, certain kinds of modulation such as phase modulation or displacement in phase space, and standard measurements such as photon counting, homodyne or heterodyne detection. The channel capacities under a mean energy constraint under these probe, modulation, and measurement choices are known in many cases [10, 21]. Using this approach, performance bounds for the communication or sensing of a Gaussian source were obtained in [10]. In addition, for lossless estimation of a uniform phase parameter, a lower bound exhibiting standard quantum limit (SQL) scaling, i.e., the behavior $\delta\Phi^2 \sim 1/N_S$, was obtained for coherent state probes, and a lower bound exhibiting H limit scaling was obtained for a quadrature-squeezed-state (or two-photon coherent state (TCS)) probe.

In this paper, we are concerned with lower bounds on the MSE for lossless and lossy phase estimation under a mean energy constraint N_S on the probe state ρ used to sense the phase. We will not consider coding over multiple instances of the phase, i.e., we set $k = n = 1$ in Fig. 1 so that $\mathbf{X} = \Phi \equiv \Phi$ and $\mathbf{Y} = \hat{\Phi} \equiv \hat{\Phi}$. This assumption of no coding is realistic in the single-parameter estimation problem considered here, though it may be relaxed in more general situations. In line with our earlier remarks, a large part of our work consists in estimating the classical capacity of the channel resulting from phase modulation of the probe state according to the value of Φ , while allowing arbitrary POVM measurements on the modulated states.

III. H LIMIT FOR LOSSLESS PHASE ESTIMATION

We consider the following general strategy for estimating a phase shift. An arbitrary pure probe state $|\psi\rangle_{IS}$ on the Hilbert space \mathcal{H}_{IS} of M ‘signal’ and M' ‘idler’ modes with average energy N_S in the signal modes is prepared [22]. The signal modes each undergo phase shifting by the unknown amount ϕ while the idler modes remain unaffected (see Fig. 2 for the $M = M' = 1$ case) so that we have the output state

$$\rho_\phi = U_\phi \rho_{IS} U_\phi^\dagger \quad (8)$$

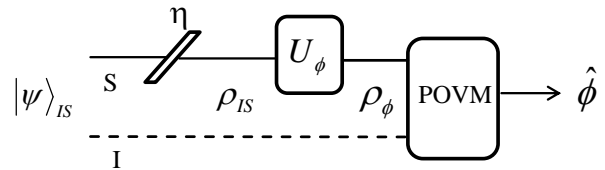


FIG. 2. Schematic of a lossless or lossy phase sensing scenario. The phase shift ϕ acts on the signal mode ‘S’. In the lossless case, we have $\eta = 1$ and the idler mode ‘I’ is not used. In the lossy case, an NDS state of the ‘S’ and ‘I’ modes is prepared and a joint measurement is made to obtain the phase estimate $\hat{\phi}$.

for

$$U_\phi = \left(\bigotimes_{m=1}^M e^{i\phi \hat{a}_m^\dagger \hat{a}_m} \right) \bigotimes I_I, \quad (9)$$

where $\{\hat{a}_m\}_{m=1}^M$ are the annihilation operators of the M signal modes and I_I is the identity operator on the M' idler modes. Finally, the optimum POVM is implemented on the joint state to yield the best MSE estimate $\hat{\Phi}$ of Φ . The above description corresponds to an entanglement-assisted parallel strategy considered in ref. [23]. It is also a special case of the general image sensing framework of ref. [23] with the number of ‘pixels’ P set to one, the number of hypotheses M , identified as the ‘number’ of different phase-shift values, going to infinity, the prior probability of the ‘image’ that shifts by ϕ set equal to $P_\Phi(\phi)$, and the cost function taken to be the mean-squared error. We will apply the results of ref. [23] extensively in the following.

In this section, we consider the lossless case in which the beam splitter of Fig. 2 has transmittance $\eta = 1$. It was shown in ref. [23] (see the section on lossless image sensing) that both multiple modes and idler entanglement are unnecessary for optimal performance in any image sensing problem in the above framework. That is, a single-mode signal-only probe state is sufficient so that we may set $M = 1$ and $M' = 0$.

The unrestricted capacity $C(N_S)$ (in nats/use) of a single-mode noiseless channel under a mean energy constraint on the channel input ensemble is well-known [24] and is given by

$$C(N_S) = (N_S + 1) \ln(N_S + 1) - N_S \ln N_S. \quad (10)$$

By ‘unrestricted’, we mean that there is no constraint on either the POVM or the modulation map other than that the output ensemble of the modulation map has mean energy N_S . This energy constraint is satisfied by the ensemble $\{P_\Phi(\phi), \rho_\phi\}$ – the $\{\rho_\phi\}$ all have the same energy, namely N_S , as the probe ρ_{IS} – so that $C(N_S) \geq C_{\text{ph}}(N_S)$, where the latter capacity is that achieved by the phase-modulated ensemble $\{P_\Phi(\phi), \rho_\phi\}$ optimized over all POVMs. Since $D(\cdot)$ of eq. (7) is a decreasing function of its argument, we may apply (7) to

get (recall that $k = n = 1$)

$$\begin{aligned}\delta\Phi^2 &\geq D(C_{\text{phase}}(N_S)) \geq D(C(N_S)) \\ &\geq \underline{D}(C(N_S)) = Q_\Phi \left(1 + \frac{1}{N_S}\right)^{-2N_S} \frac{1}{(N_S + 1)^2} \\ &\geq \frac{Q_\Phi}{e^2} \frac{1}{(N_S + 1)^2},\end{aligned}\quad (11)$$

where we have used the lower bound (6).

Eqn. (11) has the form of a H limit [5–9] with asymptotic behavior $\sim 1/N_S^2$. However, as the derivation shows, it is a non-asymptotic result valid for arbitrary prior distribution $P_\Phi(\phi)$, arbitrary probe state, arbitrary POVMs, and for all values of N_S .

It is interesting to compare the bound (11) to that obtained by Hall and Wiseman in ref. [9]. Eq. (17) of [9] reads (after removing the factor of 2 multiplying N_S as explained in [9])

$$\delta\Phi^2 \geq \frac{1}{2\pi e^3} \frac{1}{P_{\text{max}}^2(N_S + 1)^2}, \quad (12)$$

where $P_{\text{max}} \geq 1/2\pi$ is the maximum value of $P_\Phi(\phi)$. If $P_\Phi(\phi)$ is sharply peaked at some point of $[0, 2\pi)$ or not bounded from above (as when a particular ϕ has a nonzero probability of occurrence), the right-hand side of (12) can be smaller than (11), which remains nonzero as long as $P_\Phi(\phi)$ is supported on some interval of finite length. On the other hand, for Φ distributed uniformly in an interval of length L , the bound Eq. (12) coincides with eq. (11), as may be verified by setting $Q_\Phi = L^2/2\pi e$ in (11). This is rather remarkable as the two results were obtained using ostensibly quite different methods.

We may wonder if the bound (11) may be strengthened by using the capacity $C_{\text{ph}}(N_S)$ of the lossless channel restricted to phase modulation rather than the unrestricted capacity $C(N_S)$. Indeed, the ensembles well-known to achieve $C(N_S)$ are the number states with a thermal distribution [24] and coherent states with a circularly-symmetric Gaussian distribution on phase space [25]. However, for the case of uniform $P_\Phi(\phi)$, it has been recently shown that $C_{\text{ph}}(N_S)$ can be made to approach $C(N_S)$ arbitrarily closely using phase modulation on an appropriate probe state [26].

In sum, while the bounds in refs. [5–9] are applicable to more general Hamiltonians than the linear phase shift Hamiltonian considered here, we have presented a new derivation of a non-asymptotic H limit for linear optical phase estimation that is satisfied by every probe state and can deal simply with arbitrary prior information.

IV. LOWER BOUND ON THE MSE IN LOSSY PHASE ESTIMATION

We now show that the above technique can be extended to phase estimation in the presence of loss. Consider again the setup of Fig. 2 with nonzero loss so that

the beam splitter has transmittance $\eta < 1$ [27]. It follows from Theorem 1 of ref. [23] that, among all probe states $|\psi\rangle_{IS}$ with a given photon probability distribution $\{p_{\mathbf{n}}, \mathbf{n} \in \{0, 1, \dots\}^M\}$ in the M signal modes, a state of the form

$$|\Psi\rangle_{IS} = \sum_{\mathbf{n}} \sqrt{p_{\mathbf{n}}} |\Psi_{\mathbf{n}}\rangle_I |\mathbf{n}\rangle_S \quad (13)$$

minimizes the MSE, where $\{|\Psi_{\mathbf{n}}\rangle_I\}$ is any orthonormal set of idler states – such states are called NDS (Number-Diagonal Signal) states. Therefore, a lower bound on the phase estimation MSE valid for any NDS probe state of M signal modes with mean signal energy N_S is also a lower bound on any M signal-mode state (with any kind of entanglement with idler modes) of mean signal energy N_S [22]. For lossy phase estimation, unlike the lossless case of Section III, using a probe with multiple signal modes may decrease the MSE from the single mode case for the same total signal energy N_S . However, we will consider just the $M = 1$ case of a single signal-idler pair in the following.

Consider using the probe state $|\psi\rangle_{IS}$ of Eq. (13) with $M = 1$ so that $\mathbf{n} \equiv n$, an integer index. Under phase modulation with prior density $P_\Phi(\phi)$, we obtain the output ensemble $\{P_\Phi(\phi), \rho_\phi\}$ given via eq. (8), where ρ_{IS} may be written as

$$\rho_{IS} = \sum_l q_l |\chi_l\rangle_{IS} \langle\chi_l|, \quad (14)$$

where q_l is the probability that l photons are lost to the environment during the beam-splitter interaction of Fig. 2 and is given by

$$q_l = \sum_{n \geq l} p_n B_\eta(n, l) = \sum_{n \geq l} p_n \binom{n}{l} \eta^{n-l} (1 - \eta)^l. \quad (15)$$

The states $\{|\chi_l\rangle_{IS}\}$ in eq. (14) are given by

$$|\chi_l\rangle_{IS} = \frac{1}{\sqrt{q_l}} \sum_{n \geq l} \sqrt{p_n B_\eta(n, l)} |\Psi_n\rangle_I |n - l\rangle_S \quad (16)$$

and form an orthonormal set, i.e., ${}_{IS}\langle\chi_l|\chi_{l'}\rangle_{IS} = \delta_{l,l'}$ by virtue of the fact that the $\{|\Psi_n\rangle_I\}$ are orthonormal. Thus, ρ_ϕ of Fig. 2 is given by

$$\rho_\phi = \sum_l q_l |\chi_l(\phi)\rangle_{IS} \langle\chi_l(\phi)|, \quad (17)$$

where

$$\chi_l(\phi) = \frac{1}{\sqrt{q_l}} \sum_{n \geq l} \sqrt{p_n B_\eta(n, l)} e^{in\phi} |\Psi_n\rangle_I |n - l\rangle_S. \quad (18)$$

We adopt the following strategy for obtaining a lower bound on the MSE in lossy phase estimation for an NDS input. We will first obtain an upper bound $\bar{C}_{\text{ph}}(N_S)$ on the capacity of a lossy phase modulation channel restricted to an NDS input state $|\psi\rangle_{IS}$ of mean signal energy N_S . Since the lower bound Eq. (6) is decreasing

in R , the Information Transmission Inequality gives the lower bound $D(\bar{C}_{\text{ph}}(N_S))$ on the MSE. Finally, since an $M = 1$ NDS state minimizes the MSE among all states with $M = 1$, the previous bound is also valid for all $M = 1$ states with mean signal energy N_S .

In order to estimate $C_{\text{ph}}(N_S)$, we use the Holevo bound [16, 24, 28] which states

$$C_{\text{ph}}(N_S) \leq S(\bar{\rho}) - \int_0^{2\pi} d\phi P_{\Phi}(\phi) S(\rho_{\phi}), \quad (19)$$

$$= S(\bar{\rho}) - S(\rho_{IS}) \quad (20)$$

where

$$\bar{\rho} := \int_0^{2\pi} d\phi P_{\Phi}(\phi) \rho_{\phi} \quad (21)$$

is defined to be the average output density operator, and $S(\cdot)$ denotes von Neumann entropy. Defining the uniform signal-phase randomization CP map \mathcal{P} on \mathcal{H}_{IS} as

$$\mathcal{P}\sigma = \int_0^{2\pi} \frac{d\phi}{2\pi} U_{\phi} \sigma U_{\phi}^{\dagger}, \quad \sigma \in \mathcal{H}_{IS}, \quad (22)$$

we have $S(\bar{\rho}) \leq S(\mathcal{P}\bar{\rho})$ because \mathcal{P} is a unital CP map [16]. It is readily verified that $\mathcal{P}\bar{\rho}$ is independent of $P_{\Phi}(\phi)$ and equals

$$\mathcal{P}\bar{\rho} = \sum_{l=0}^{\infty} \sum_{n=l}^{\infty} p_n B_{\eta}(n, l) |\Psi_n\rangle_I \langle \Psi_n| \otimes |n-l\rangle_S \langle n-l|. \quad (23)$$

The orthogonality of $\{|\Psi_n\rangle_I\}$ then implies that

$$S(\mathcal{P}\bar{\rho}) = H(N, N-L), \quad (24)$$

where N and $N-L$ are the classical random variables corresponding to a measurement on $\mathcal{P}\bar{\rho}$ of the $\{|\Psi_n\rangle_I\}$ basis on the idler mode and the photon number in the signal mode respectively, and $H(\cdot)$ is the Shannon entropy. Similarly, the orthogonality of $\{|\chi_l\rangle_{IS}\}$ implies that

$$S(\rho_{IS}) = H(L). \quad (25)$$

Combining the above facts, we have

$$C_{\text{ph}}(N_S) \leq S(\mathcal{P}\bar{\rho}) - S(\rho_{IS}) \quad (26)$$

$$\begin{aligned} &= H(N, N-L) - H(L) \\ &= H(N, L) - H(L) \\ &= H(L|N) - [H(L) - H(N)] \\ &= \sum_n p_n H(L|N=n) - [H(L) - H(N)] \end{aligned} \quad (27)$$

$$\begin{aligned} &\leq \sum_n \frac{p_n}{2} \ln \left[2\pi e \left(\eta(1-\eta)n + \frac{1}{12} \right) \right] \\ &\quad - [H(L) - H(N)] \end{aligned} \quad (28)$$

$$\begin{aligned} &\leq \frac{1}{2} \ln \left[2\pi e \left(\eta(1-\eta)N_S + \frac{1}{12} \right) \right] \\ &\quad - [H(L) - H(N)]. \end{aligned} \quad (29)$$

FIG. 3. The single-mode input state and loss channel \mathcal{L} for which the entropy gain from input to output equals $H(L) - H(N)$ of Eq. (29).

Here we have used standard entropy manipulations to obtain eq. (27). To obtain (28), we have applied the bound

$$H(X) \leq \frac{1}{2} \ln \left[2\pi e \left(\text{Var } X + \frac{1}{12} \right) \right] \quad (30)$$

on the Shannon entropy $H(X)$ of a discrete random variable X in terms of its variance (see Problem 8.7, p. 258 of [13]), substituting the variance of L conditioned on $N = n$ (L has a binomial distribution in this case). Finally, eq. (29) follows from concavity of the logarithm.

We now bound the second term in eq. (29). Consider the single-mode pure loss channel \mathcal{L} depicted in Fig. 3. It is readily verified that, for the input state $\rho_{\text{in}} = \sum_n p_n |n\rangle_S \langle n|$, the channel outputs the state $\rho_{\text{out}} = \sum_l q_l |l\rangle_S \langle l|$ so that the *entropy gain* from input to output is precisely $H(L) - H(N)$. For the channel \mathcal{L} of Fig. 3, Holevo has shown (See Theorem 2 of [29]) that the *minimum* entropy gain

$$\inf_{\rho_{\text{in}} \in \mathcal{H}_S} [S(\mathcal{L}\rho_{\text{in}}) - S(\rho_{\text{in}})] = \ln(1-\eta), \quad (31)$$

where the infimum is over all input states in \mathcal{H}_{IS} and therefore includes the input state of Fig. 3. We thus have, for all probe states $|\psi\rangle_{IS}$,

$$\ln(1-\eta) \leq H(L) - H(N), \quad (32)$$

which, combined with (29), gives

$$\begin{aligned} C_{\text{ph}}(N_S) &\leq \frac{1}{2} \ln \left[\frac{2\pi e}{(1-\eta)^2} \left(\eta(1-\eta)N_S + \frac{1}{12} \right) \right] \\ &\equiv \bar{C}_{\text{ph}}(N_S), \end{aligned} \quad (33)$$

which is the sought upper bound on $C_{\text{ph}}(N_S)$.

The final step is to use the lower bound (6) and the Information Transmission Inequality to get the lower bound

$$\begin{aligned} \delta\Phi^2 &\geq \underline{D}(\bar{C}_{\text{ph}}(N_S)) \\ &= \frac{Q_{\Phi}(1-\eta)^2}{2\pi e [\eta(1-\eta)N_S + \frac{1}{12}]} \end{aligned} \quad (34)$$

on the MSE that exhibits SQL scaling $\delta\Phi^2 \sim 1/N_S$ for large N_S , implying that for any nonzero amount of loss,

nonclassical states of light are not much superior to coherent states for phase sensing, at least when $M = 1$. Note that the bound (34) does not reduce to Eq. (11) as $\eta \rightarrow 1$ because the estimate (32) is not tight in that limit [30]. The following bound applicable to lossy phase estimation that is in terms of the mean and variance of the input state and exhibits SQL scaling has been obtained in [31]:-

$$\delta\Phi_\phi^2 \geq \frac{1-\eta}{4\eta N_S} + \frac{1}{4\langle\Delta\hat{N}_S^2\rangle}. \quad (35)$$

The left-hand side $\delta\Phi_\phi^2$ is the mean-squared error achieved for a particular (but arbitrary) value of ϕ and $\langle\Delta\hat{N}_S^2\rangle$ is the variance of the signal photon number. This bound, based as it is on the quantum Cramér-Rao inequality, is valid for every value of ϕ provided the estimate $\hat{\Phi}$ is unbiased. Note that, in the region of large $\langle\Delta\hat{N}_S^2\rangle$, which is achievable for any finite N_S [32], the bounds (35) and (34) are rather similar in form.

V. CONCLUSION

We have developed the rate-distortion approach to lower bounds on the MSE of optical phase estimation.

We have obtained an H limit for lossless phase estimation and a lower bound with SQL scaling in the presence of loss. It is hoped that the approach of this work can be extended to obtain performance bounds on the estimation of other system parameters using quantum states of light.

VI. ACKNOWLEDGMENTS

Valuable discussions with Mankei Tsang and Brent J. Yen are much appreciated. I also thank Horace P. Yuen for introducing me to the rate-distortion approach in quantum metrology via his article [10]. This material is based on work supported by the Singapore National Research Foundation under NRF Grant No. NRF-NRFF2011-07.

-
- [1] B. Yurke, S. L. McCall, J. R. Klauder, Phys. Rev. A **33**, 4033 (1986); B. C. Sanders, G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995); Z. Y. Ou, Phys. Rev. Lett. **77**, 2352 (1996).
 - [2] P. M. Anisimov, et al., Phys. Rev. Lett. **104**, 103602 (2010).
 - [3] A. Rivas and A. Luis, arXiv:1105.6310v2 (2011).
 - [4] Y. R. Zhang, et al., arXiv:1105.2990v2 (2011).
 - [5] V. Giovannetti, S. Lloyd, and L. Maccone, arXiv:1109.5661 (2011).
 - [6] M. J. W. Hall, D. W. Berry, M. Zwierz, and H. M. Wiseman, arXiv:1111.0788v2 (2012).
 - [7] M. Tsang, arXiv:1111.3568v3 (2011).
 - [8] V. Giovannetti and L. Maccone, arXiv:1201.1878 (2012).
 - [9] M. J. W. Hall and H. M. Wiseman, New J. Phys. **14**, 033040 (2012).
 - [10] H. P. Yuen, in *Quantum Squeezing* (eds. P. D. Drummond and Z. Ficek, Springer Verlag 2004, Chap. 7); arXiv:quant-ph/0109054 (2001).
 - [11] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression* (Englewood Cliffs, NJ: Prentice-Hall, 1971).
 - [12] C. E. Shannon, Bell Sys. Tech. Jnl. **27**, pp. 379-423, pp. 623-656 (1948).
 - [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 2nd Ed, 2006).
 - [14] C. E. Shannon, IRE International Convention Record, **7**, 325 (1959).
 - [15] T. Berger and J. D. Gibson, IEEE Trans. on Info. Theory **44**, 2693 (1998).
 - [16] M. Nielsen, and I. Chuang, *Quantum Information and Computation* (Cambridge Univ. Press, Cambridge, 2000).
 - [17] M. M. Wilde, *From Classical to Quantum Shannon Theory*, arXiv:1106.1445 (2011).
 - [18] B. Schumacher, Phys. Rev. A **51**, 2738 (1996).
 - [19] H. Barnum, Phys. Rev. A **62** 042309 (2000); I. Devetak and T. Berger, IEEE Trans. on Inform. Th. **48**, 1580 (2002); X.-Y. Chen and W.-M. Wang, IEEE Trans. on Inform. Th. **54**, 743 (2008); N. Datta, M.-H. Hsieh, and M. M. Wilde, arXiv:11084940 (2011).
 - [20] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part II* (John Wiley & Sons, New York, 1971).
 - [21] J. H. Shapiro, IEEE J. Sel. Top. Quant. Elec. **15**, 1547 (2009).
 - [22] It was shown in [23] that mixed probe states are not superior to pure ones for lossless or lossy image sensing.
 - [23] R. Nair and B. J. Yen, Phys. Rev. Lett. **107**, 193602 (2011).
 - [24] H. P. Yuen and M. Ozawa, Phys. Rev. Lett. **70**, 363 (1993).
 - [25] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, Phys. Rev. Lett. **92**, 027902 (2004).
 - [26] S. Guha, Z. Dutton, R. Nair, J. H. Shapiro, and B. J. Yen, in *Frontiers in Optics/Laser Science*, paper LTuF2, OSA Technical Digest (Optical Society of America, 2011).
 - [27] We note that, in Fig. 2, the output state ρ_ϕ is independent of the order in which the loss interaction and phase shift are applied.
 - [28] A. S. Holevo, Prob. Info. Trans. **9**, 177 (1973).

- [29] A. S. Holevo, Dokl. Math. **82**, 730 (2010); arXiv:1003.5765v1 (2010).
- [30] It is possible that a lower bound on the entropy gain of the loss channel that takes into account an energy constraint of N_S on the input state may lead to an improved estimate.
- [31] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nature Phys. **7**, 406 (2011).
- [32] For any given $N_S > 0$, $\langle \Delta \hat{N}_S^2 \rangle$ can be made arbitrarily large.